Behavior-driven forecasts of neighborhood-level COVID-19 spread in New York City

Authors

Renquan Zhang,[1] Jilei Tai,[1] Qing Yao,[2] Wan Yang,[3,4] Kai Ruggeri,[5] Jeffrey Shaman,[2,6] Sen Pei[2]*

Affiliations

[1]School of Mathematical Sciences, Dalian University of Technology, Dalian, 116024, China

[2]Department of Environmental Health Sciences, Mailman School of Public Health, Columbia University, New York, NY 10032

[3]Department of Epidemiology, Mailman School of Public Health, Columbia University, New York, NY 10032

[4]Herbert Irving Comprehensive Cancer Center, Columbia University Medical Center, New York, NY 10032

[5]Department of Health Policy and Management, Mailman School of Public Health, Columbia University, New York, NY 10032

[6]Columbia Climate School, Columbia University, New York, NY 10027

*Corresponding author. Email: sp3449@cumc.columbia.edu

Abstract

The COVID-19 pandemic in New York City (NYC) was characterized by marked disparities in disease burdens across neighborhoods. Accurate neighborhood-level forecasts are critical for planning more equitable resource allocation; however, such spatially high-resolution forecasts remain scarce in operational use. Here, we analyze aggregated foot traffic data derived from mobile devices to measure the connectivity among 42 NYC neighborhoods driven by various human activities such as dining, shopping, and entertainment. Using real-world time-varying contact patterns in different place categories, we develop a parsimonious behavior-driven epidemic model that incorporates population mixing, indoor crowdedness, dwell time, and seasonality of virus transmissibility. This process-based model supports accurate modeling of neighborhood-level SARS-CoV-2 transmission throughout 2020. In the best-fitting model, we estimate that the force of infection in indoor settings increases sublinearly with crowdedness and dwell time. Retrospective forecasting demonstrates that this behavior-driven model generates improved short-term forecasts in NYC neighborhoods. This model may be adapted for use with other respiratory pathogens sharing similar transmission routes.

Introduction

In urban settings, COVID-19 has imposed differential disease and healthcare burdens across neighborhoods characterized by varying demographic and socioeconomic conditions (*1*–*3*). A case in point is New York City (NYC), where pronounced health disparities within the city have been reported (*4*–*7*). The stark difference in disease burdens requires the delivery of tailored services and support to local communities. Accurate forecasting of respiratory disease outbreaks at the neighborhood level can inform policymaking to design more equitable interventions and resource allocation in future epidemics. Nonetheless, generating such spatially high-resolution forecasts remain challenging due to the significant behavioral variability and interconnectedness among neighborhoods in metropolitan areas.

The transmission dynamics of SARS-CoV-2 were collectively shaped by human behavior, seasonal variation in virus transmissibility (i.e., seasonality), population immunity, and pathogen evolution (*8*). Among those factors, human behaviors, both voluntary and those driven by governmental interventions, played a major role during the early phase of the pandemic (*9*–*15*). However, early studies examining the associations between human mobility indices (e.g., distance traveled, time spent at home, or percentage change in mobility) and COVID-19 spread yielded inconclusive results – strong associations during lockdowns (*16*–*18*) but weak or no associations during the summer of 2020 (*19*, *20*). A potential explanation for this inconsistency is that the mobility indices employed did not accurately reflect behavioral effects on the contagion process. For instance, a person can drive a long distance alone without contributing to disease transmission, yet a short visit to a convenience store may involve multiple exposures. As a result, incorporating mobility indices into process-based models may not necessarily improve disease forecasting.

Foot traffic data derived from mobile devices offer an effective means of tracking human mobility and behavior patterns. Unlike aggregated mobility indices, foot traffic data document visitation records at points of interest (POIs) with high spatial and temporal resolution, providing a direct measurement of population contacts in various locations. These privacy-preserving data have supported studies on high-resolution disease modeling (*21*–*25*), risk of airborne transmission (*26*), and human mobility changes during outbreaks or natural disasters (*27*, *28*). For instance, mobility network models and agent-based models were used to simulate epidemic spread within small geographical units, which provided valuable insights on inequalities of infection risk across communities (*21*–*25*). However, most of these models were calibrated against aggregated disease data at the county or city level. It remains unclear whether they can reproduce heterogeneous SARS-CoV-2 transmission at the neighborhood level and support forecasting with validated skills. In addition, the complexity of these models makes it challenging for rapid deployment by public health officials during an emergency.

In this study, we aim to develop a parsimonious process-based model informed by aggregated foot traffic data to generate validated neighborhood-level forecasts with efficient simulation and calibration. We analyzed foot traffic data within various place categories in NYC during 2020. Densely populated metropolitan areas like NYC often experience earlier outbreaks, which may then propagate rapidly to other locations. During the COVID-19 pandemic, NYC was the first epicenter in the US. Using fine-grained foot traffic data, we identified distinct mobility and indoor contact patterns across five place categories: restaurants and bars, retail, arts and entertainment, educational settings, and other places. We then developed a behavior-driven epidemic model informed by category-

specific mobility and indoor contact pattern data. Our results indicate that this model, capturing social drivers of contagion (i.e., human contacts in various activities) and transmission seasonality, can fit COVID-19 spread across the 42 NYC neighborhoods during 2020. Neighborhood-level model fitting suggests that the force of infection in these place categories increases sublinearly with crowdedness and dwell time. Furthermore, we demonstrated that this model, in conjunction with data assimilation techniques, produced more accurate short-term forecasts of COVID-19 spread at the neighborhood level during the first year of the pandemic.

Results

Mobility and contact patterns driven by human activities

Human contacts are driven by various activities such as dining, shopping, work, and entertainment. We used foot traffic data shared by SafeGraph (*29*) to track mobility and contact patterns during these activities (Materials and Methods). Using the North American Industry Classification System (NAICS) code (*30*), we classified all POIs into five categories: (1) restaurants and bars, (2) retail, (3) arts and entertainment, (4) educational settings, and (5) other places (Table S1). A clustering analysis indicates that POIs within each category exhibit similar crowdedness, dwell time, and variance of visitor numbers during a week (Fig. S1). This categorical classification allows parsimonious representation of population mixing in different settings and aligns with the practical implementation of governmental interventions targeting specific high-risk sectors.

The foot traffic data recorded the overall temporal trend of mobility in NYC. The daily number of visitors to all place categories plummeted in March 2020 and was reduced by 81% at its lowest point on April 12, 2020 (Fig. 1A), following strict control measures enforced during the initial wave. Thereafter, mobility gradually increased throughout 2020 but did not recover to its pre-pandemic level.

The foot traffic data provided insights on the distribution of built environments within NYC. The NYC Department of Health and Mental Hygiene (DOHMH) uses United Hospital Fund (UHF) Areas as one way to designate neighborhoods (Fig. 1B). As the spatial allocation of POIs can significantly influence the mobility of residents engaged in various activities, we examined the number of POIs located within each neighborhood. We observed heterogeneous geographical distributions of POIs and distinct spatial patterns across place categories (Fig. S2), revealing spatial disparities in urban services and facilities. For instance, restaurants and bars concentrated in lower Manhattan and drew visitors from distant neighborhoods (Fig. 1B), whereas educational facilities exhibited a more even distribution across neighborhoods (Fig. S2).

The foot traffic data captured the interconnectedness between neighborhoods driven by different human activities. To assess population mixing between the 42 UHF neighborhoods, we constructed daily mobility matrices for each place category, denoted by $\{C_{ij}^{p}(t)\}$, where $C_{ij}^{p}(t)$ represents the number of visitors from home neighborhood i to destination neighborhood j in place category p on day t. We visualized average daily mobility for four place categories (excluding other places that lump in different POI types) from January 6, 2020, to March 1, 2020 (Fig. 2), depicting mobility patterns in winter prior to the pandemic. Notably, regardless of place categories, most visitors went to POIs in their home neighborhoods, as evidenced by the diagonal elements with large daily visitor counts. A clear cross-neighborhood mobility structure emerged within four

boroughs (the Bronx, Brooklyn, Queens, and Staten Island), indicating that residents in these boroughs frequently visited local POIs close to their residence. In contrast, Manhattan, the primary employment borough with high living costs (*31*), drew large numbers of visitors from other boroughs. Cross-neighborhood mobility was most pronounced for restaurants and bars and least notable for educational facilities. We compiled daily mobility matrices for 2020 to track the temporal evolution of cross-neighborhood population mixing.

The foot traffic data also revealed contact patterns within POIs that may modulate infection risk in indoor environments. Particularly, we focused on the crowdedness (measured by daily number of visitors per square meter) (*32*) and average dwell time of visitors (Materials and Methods) (*21*). Among the five examined place categories, restaurants and bars had on average the highest crowdedness (Fig. 1C). The crowdedness in all place categories decreased by up to 82% (arts and entertainment) following the lockdown in March 2020 and then moderately increased in the latter half of 2020. In contrast, dwell time in all place categories remained relatively stable (Fig. 1D), except for a temporary increase during the lockdown and an increase in educational facilities after the start of the fall semester in September. Variations in crowdedness and dwell time may lead to differential infection risk in POIs.

A behavior-driven epidemic model

To model contagion processes driven by human activities, we developed a behavior-driven epidemic model informed by real-world mobility matrices for different place categories, time-varying crowdedness and dwell time, and seasonality of virus transmissibility (Materials and Methods). Denote $\lambda_j^p(t)$ as the force of infection (FOI) in place category p in neighborhood j on day t. We linked FOI to crowdedness and dwell time through:

$$\lambda_j^p(t) = \frac{v_p \left(d_j^p(t)\right)^a \left(\tau_j^p(t)\right)^b I_j^p(t)}{N_j^p(t)}, \quad [1]$$

where $d_j^p(t)$ and $\tau_j^p(t)$ are the crowdedness and average dwell time in place category p in neighborhood j on day t. $I_j^p(t)$ and $N_j^p(t)$ denote the infectious and total population present in place category p in neighborhood j on day t. The parameters a and b determine the nonlinear relationship between virus transmissibility and crowdedness and dwell time, respectively. Note, a and b are shared by all place categories, representing the universal functional form linking crowdedness and dwell time to FOI that applies in all POIs. In practice, this functional form links real-world intervention criteria such as occupancy limit and operation hours, which are straightforward to measure and control, to the transmission risk of infectious diseases. We further introduced a multiplier, v_p, for each place category representing the compound effects of potential sampling bias across different place categories, category-specific interventions, relative contribution to infection risk in different place categories, and other unaccounted factors.

In addition to population mixing, the transmissibility of SARS-CoV-2 exhibits strong seasonality modulated by meteorological factors, with humidity playing the major role in the northern US (*33*). We imposed a seasonality term, $\omega(t)$, forced by absolute humidity

(AH), inspired by the modeling of influenza seasonality (*34*, *35*). We further modeled the time-varying ascertainment rate (*36*) (Fig. S3) and delay from infection to case confirmation (Materials and Methods).

Estimating the dependency of FOIs on crowdedness and dwell time

We first used grid search and model selection to estimate the parameters a and b that control the impacts of crowdedness and dwell time on FOIs. We tested 21×21 parameter combinations: $a \in [0,2]$ and $b \in [0,2]$ with a step of 0.1 for both parameters (including $a = 0$ and $b = 0$). For each parameter combination, we fixed a and b in the model and calibrated the model against real-world weekly COVID-19 cases in 42 neighborhoods from March 1, 2020 to June 7, 2020, a period with substantial variation in crowdedness and dwell time (Fig. 1). Model calibration was performed using Markov Chain Monte Carlo (MCMC) (Materials and Methods) (*37*). The goodness-of-fit for each parameter combination was measured using loglikelihood, averaged over 100 samples in the MCMC chain (selected every 10 steps from the last 1,000 samples) (Materials and Methods). The landscape of loglikelihood indicates that the best-fitting model was roughly in the region of $a \in [0,1]$ and $b \in [0,1.5]$ (Fig. 3A).

To more accurately estimate the parameters a and b in the best-fitting model, we examined the profile loglikelihood for both parameters. Profile likelihood method is widely used to estimate confidence intervals (CIs) of parameters in epidemic models (*38*, *39*). We fit the loglikelihood surface using a cubic spline interpolation and identified the maximum loglikelihood at $a = 0.338$ and $b = 0.270$ on the surface. To further quantify the uncertainty of a, we plotted the profile loglikelihood curve by setting $b = 0.270$ in the surface function (Fig. 3B). The 95% CI of a was estimated to be $[0.285, 0.464]$ (Materials and Methods). Similarly, we estimated the 95% CI of b to be $[0.156, 0.342]$ (Fig. 3C). This model selection result suggests that the best-fitting model supports that FOIs increases sublinearly with crowdedness and dwell time. In other words, disease transmission risk increases rapidly when crowdedness and dwell time rise from lower values; however, the increase rate declines for larger crowdedness and longer dwell time, showing a diminishing-return effect. In following analyses, we fixed $a = 0.338$ and $b = 0.270$ in the behavior-driven epidemic model. The posterior model parameters estimated using this setting are reported in Table S3.

Modeling neighborhood-level COVID-19 spread

We fit the behavior-driven epidemic model to weekly COVID-19 cases in 42 neighborhoods from March 1, 2020 to December 13, 2020, before the availability of COVID-19 vaccines in NYC. In model fitting, we fixed epidemiological and seasonality parameters and estimated v_p for five place categories and β_h for baseline transmission. As model parameters likely changed over the course of 2020 due to shifting interventions and behaviors, we estimated the parameters separately for five distinct periods, defined by the dates with major policy changes in NYC (Fig. 4A, Supplementary Materials, Table S2). We estimated posterior parameters using MCMC and confirmed that MCMC chains converged using diagnostic statistics (Table S4). We then ran model simulations using estimated parameters and compared the simulated COVID-19 case counts with reported numbers from the 42 neighborhoods.

At the city level, the behavior-driven epidemic model reproduced the trend of COVID-19 cases throughout 2020 (Fig. 4A). In addition, model simulations agreed with observations in most neighborhoods (Fig. S4), suggesting that aggregated foot traffic data combined with seasonality can support neighborhood-level modeling of SARS-CoV-2 transmission. Model fitting in several neighborhoods such as Greenwich Village-SoHo had larger discrepancy, possibly due to model misspecification or issues in the mobility data. Examples in four neighborhoods in the Bronx, Brooklyn, Manhattan, and Queens are provided in Fig. 4B.

Retrospective forecasts of neighborhood-level COVID-19 cases

While the behavior-driven epidemic model can reproduce real-world neighborhood-level SARS-CoV-2 transmission, it is unknown if this model can generate more accurate short-term forecasts of COVID-19 cases. We coupled the behavior-driven epidemic model with an efficient data assimilation algorithm, the ensemble adjustment Kalman filter (EAKF) (*40*), to generate retrospective forecasts at the neighborhood level (Material and Methods, Supplementary Materials). During the forecasting process, model parameters and variables were updated weekly once real-world COVID-19 surveillance data became available. We then integrated the optimized model into the future to generate forecasts. Within the forecasting horizon, mobility matrices, indoor contact patterns, and reporting rates were fixed using the information available when the forecasts were generated.

The one-week ahead forecasts of COVID-19 cases generally agreed with observations at the neighborhood level (Fig. S5), capturing the large variation in disease burdens. We compared the forecast skill of the behavior-driven epidemic model with a baseline metapopulation model without the place category-specific mobility and FOIs informed by crowdedness and dwell time (Materials and Methods). The baseline model used the same model-data assimilation framework and incorporated mobility data across neighborhoods obtained from the same foot traffic data. A similar framework has been used to model COVID-19 spread at the US county level (*41*, *42*). The baseline model configurations were tuned to ensure decent forecasting performance at the neighborhood level (Fig. S6).

The behavior-driven epidemic model outperformed the baseline model in most neighborhoods and weeks. We compared the one-week ahead forecasts generated for each neighborhood at each week using three metrics: (1) mean absolute percentage error (MAPE) for point predictions, and (2) log score and (3) weighted interval score (WIS) (*43*) for probabilistic predictions (Materials and Methods). Log score and WIS measure the uncertainty of predictive distributions and have been used as the standard metrics to evaluate real-time forecasts for influenza (*44*) and COVID-19 (*45*). Specifically, lower MAPEs, higher log scores, and lower WIS scores indicate better forecasts. In general, the proposed model generated more accurate forecasts in most neighborhoods and forecast weeks (Fig. 5). The comparison with the baseline model suggests that the inclusion of place category-specific FOIs linked to crowdedness and dwell time likely improved forecasts. The behavior-driven epidemic model also generated improved two-week ahead forecasts (Figs. S7-S9).

Discussion

Population contacts driven by human activities can facilitate the transmission of SARS-CoV-2. In this study, we analyzed aggregated foot traffic data in different place categories

in NYC. Our study resulted in several findings that can inform improved neighborhood-level forecasting of disease outbreaks.

First, we identified differential mobility and indoor contact patterns in different place categories, likely influenced by the geographical distribution of POIs, the nature of human activities (e.g., dining versus shopping), and socioeconomic factors (e.g., working from home versus in an essential business) (*46*, *47*). These differences explained heterogeneous disease burdens across NYC neighborhoods in our process-based model.

Second, aggregated foot traffic data for typical human activities are sufficient to support neighborhood-level infectious disease modeling without privacy concerns. A fundamental question in infectious disease modeling is whether the inclusion of more detailed processes results in more precise epidemic simulation, and to what extent system granularity is needed to inform real-world application of model outcomes. High-granularity epidemic models informed by human contact data are increasingly used in real-time forecasting. However, fine-grained contact data impose strong structural constraints on model dynamics, creating challenges in calibrating high-dimensional models to high-resolution real-world data (*48*). The behavior-driven epidemic model, working at the neighborhood scale, provides a tradeoff between model realism, privacy protection, and computational efficiency.

Third, the inclusion of place category-specific FOIs in the behavior-driven epidemic model improved short-term forecasting at the neighborhood level. The more precise representation of infection risk in different place categories enhanced the forecast skills of the model. However, the proposed forecasting system, which assumes constant mobility and contact patterns in the forecast horizon, may not improve longer-term forecasts. This limitation underscores the necessity of predicting behavior changes in response to disease outbreaks and modeling the feedback loop between behaviors and epidemics. While theoretical work exists on this topic (*49*–*56*), empirical evidence of the impacts of policies and behaviors on epidemic spread needs to be quantified more precisely (*57*). Currently, data-driven modeling applications and validations in real-world settings remain scarce.

While the behavior-driven epidemic model reproduced COVID-19 spread across NYC neighborhoods in 2020, process-based modeling beyond 2020 must additionally consider population immunity and virus evolution. Cumulative infections, vaccination, and immunity waning collectively shaped the immunological landscape upon which SARS-CoV-2 circulated after 2020 (*58*). The continuous evolution of SARS-CoV-2 led to the emergence of new variants with varying transmissibility, immune escape capability, and disease severity (*59*, *60*). Modeling these complex processes poses a significant challenge, especially at local scales that hold direct relevance to policymaking.

The behavior-driven epidemic model, formulated in a metapopulation structure, can be deployed in real time to support outbreak response. As mobility and indoor contact patterns can concomitantly modulate the transmission of a range of respiratory pathogens, this model may be generalized to work for other respiratory pathogens sharing similar transmission routes with proper modifications.

Materials and Methods

Data

Real-world weekly COVID-19 confirmed cases aggregated in 42 UHF neighborhoods were provided by the NYC DOHMH to support model fitting and retrospective forecasting. Case data in congregated settings were excluded to represent community transmission of SARS-CoV-2. We tracked mobility and contact patterns of NYC residents using aggregated foot traffic data covering 90,164 POIs. For each POI, the hourly number of visitors and home locations of visitors down to the census block groups level were recorded. To enhance privacy, SafeGraph excludes census block group information if fewer than five devices visited an establishment in a month from a given census block group. Other relevant metadata of POIs include the type of place (NAICS code), physical area (square meters), and daily average dwell time of visitors (minutes). We constructed the daily mobility matrices for each place category across the 42 neighborhoods by aggregating visitors to relevant POIs at the neighborhood level. For visitors without home information, we assigned their home neighborhoods using the distribution of visitors with home information in the same POI. Crowdedness for each POI was measured by dividing the daily visitor count by its physical area. Daily crowdedness and dwell time for each place category were averaged across all POIs within that category. Daily absolute humidity for NYC was derived from North American Land Data Assimilation System data (*61*).

The behavior-driven epidemic model

The transmission dynamics are described by the following equations.

$$\frac{dS_i}{dt} = -S_i \sum_{j,p} M_{ij}^p(t)\lambda_j^p(t)\omega(t) - \frac{\beta_h \omega(t) S_i I_i}{N_i}, \quad [2]$$

$$\frac{dE_i}{dt} = S_i \sum_{j,p} M_{ij}^p(t)\lambda_j^p(t)\omega(t) + \frac{\beta_h \omega(t) S_i I_i}{N_i} - \frac{E_i}{Z}, \quad [3]$$

$$\frac{dI_i}{dt} = \frac{E_i}{Z} - \frac{I_i}{D}. \quad [4]$$

Here N_i, S_i, E_i, and I_i are the total, susceptible, exposed, and infectious population in neighborhood i; Z and D are the latency and infectious duration; β_h is the baseline transmission rate in places not captured by the foot traffic data; $M_{ij}^p(t)$ represents the fraction of population living in neighborhood i visiting place category p in neighborhood j on day t; $\lambda_j^p(t)$ is the force of infection (FOI) in place category p in neighborhood j on day t, parameterized using crowdedness and dwell time; and $\omega(t)$ imposes seasonality on the transmissibility of SARS-CoV-2. FOIs are defined in Eq. [1]. We define $\omega(t) \propto w_{min} + \exp(-180 \times q(t) + \log(w_{max} - w_{min}))$, where w_{max} and w_{min} are the parameters controlling the intensity of seasonality, and $q(t)$ is the time-varying specific humidity, a measure of absolute humidity. This functional form was inspired by the seasonality of influenza, which has been validated by influenza modeling and forecasting (*34*, *35*). We normalized $\omega(t)$ so that the average value over a year is one. For the first year of the pandemic, we did not consider immunity loss as studies showed re-infections were rare (*62*). Note that an individual may be counted as a visitor to multiple place categories on the same day; however, the relative infection risk in multiple place categories on the same day (e.g., modulated by the time spending in different place categories) can be factored in the parameter v_p in FOIs and estimated through model

fitting. As a result, this model structure addresses the issue of double-counting susceptible population in each neighborhood. The model was integrated daily using the Euler stepping scheme. To introduce stochasticity, we generated new infections using a Poisson distribution with mean values computed in Eqs. [2-3].

Model configuration and observation model

The behavior-driven epidemic model was initiated on February 26, 2020, prior to the reporting of the first COVID-19 case in NYC on March 1, 2020. Initial infections on the first day of simulation were informed by estimated infections that accounted for the severe underreporting during the early phase of the pandemic (*63*). Specifically, we randomly drew initial infection numbers in NYC from a uniform distribution $U[1000,2000]$ and distributed them to the 42 neighborhoods as infectious population (I) in the model according to the distribution of cumulative cases reported within three weeks after March 1, 2020. The exposed population (E) in each neighborhood was set as $E = 5I$, representing rapid increase of infections before the detection of local cases. The remainder of the population was set as susceptible (S). Epidemiological parameters (Z and D) were fixed as $Z = 3.59$ days and $D = 3.56$ days. The incubation and infectious periods were estimated using early infection data from the US (*64*). The seasonality parameters were set as $w_{max} = 2.6$ and $w_{min} = 1.4$, informed by the rough estimate of the basic reproductive number of SARS-CoV-2 in winter and summer. The seasonality of virus transmissibility was modulated by the daily absolute humidity data in NYC. Sensitivity analysis was performed for other values of w_{max} and w_{min} ($w_{max} \in \{2.5, 2.6, 2.7, 2.8, 2.9\}$ and $w_{min} \in \{1.1, 1.2, 1.3, 1.4, 1.5\}$) and the performance of fitting was robust.

Daily confirmed cases were computed using an observation model. To account for varying surveillance efforts across communities and over time, we used the weekly ascertainment rates estimated by Yang et al. (*36*) for each neighborhood validated by case, hospitalization, death, serological, and wastewater data (Fig. S3). We further introduced a gamma distribution $\Gamma(1.85, 7.57)$ to model the delay from infection to case conformation, similar to our prior works (*9*, *41*, *42*). Among the newly exposed population $E_{new}(t)$ on a given day t, α fraction will be reported. These new infections will be confirmed on day $t + t_d$, where t_d is drawn from the gamma distribution.

Markov Chain Monte Carlo and model fitting

We fit the model to weekly confirmed cases in 42 neighborhoods using Metropolis-Hastings MCMC. For each combination of parameters a and b, we performed MCMC independently. Within each MCMC fitting, v_p for five place categories and β_h for baseline transmission were estimated and other epidemiological and seasonality parameters were fixed as described in model configuration. The goal of MCMC is to estimate the distributions of model parameters $\boldsymbol{x} = (v_1, v_2, v_3, v_4, v_5, \beta_h)^T$ given observed case data $\boldsymbol{O}$: $P(\boldsymbol{x}|\boldsymbol{O}) \propto P(\boldsymbol{x})P(\boldsymbol{O}|\boldsymbol{x})$, where $P(\boldsymbol{x})$ is the prior distribution and $P(\boldsymbol{O}|\boldsymbol{x})$ is the likelihood of observing $\boldsymbol{O}$ given model parameters $\boldsymbol{x}$. The Metropolis-Hasting algorithm proceeds as follows:

1. Initialization. Select initial states of unknown parameters $\boldsymbol{x_0} = \{x_i^0\}$ from prior distributions, where x_i^0 is the initial value of the i^{th} parameter.
2. Iteration. For each parameter x_i^t at iteration t, sequentially perform the following procedures while fixing other parameters at their current states.

a. Generate a random candidate state x_i' according to a proposal distribution $g(x_i'|x_i^t)$.
b. Calculate the acceptance probability $A(x_i', x_i^t) = \min\left(1, \frac{P(\boldsymbol{x}'|\boldsymbol{O})}{P(\boldsymbol{x_t}|\boldsymbol{O})}\frac{g(x_i^t|x_o')}{g(x_i'|x_i^t)}\right)$, where $\boldsymbol{x_t}$ is the current model parameter vector and $\boldsymbol{x}'$ is the perturbed vector with x_i^t replaced by x_i'.
c. Accept or reject the candidate state x_i'. Draw a uniform random number $u \in U[0,1]$. If $u \leq A(x_i', x_i^t)$, update x_i^t to x_i'; otherwise, keep x_i^t unchanged.
d. After looping through all parameters, move to the next iteration $t = t + 1$.

In MCMC, the prior distributions of parameters were set broadly as $v_p \in U[0,100]$ and $\beta_h \in U[0,1]$. The ranges of these parameters were explored and selected to ensure that the model can produce outbreaks with a similar scale as in the real-world data. The proposal distributions were selected as symmetric. For v_p, we perturbed the current parameter value by adding a Gaussian random noise with a standard deviation of 0.5: $g(x_i'|x_i^t) = \mathcal{N}(\mu = x_i^t, \sigma = 0.5)$; for β_h, we used a Gaussian random noise with a standard deviation of 0.01: $g(x_i'|x_i^t) = \mathcal{N}(\mu = x_i^t, \sigma = 0.01)$. We next compute the acceptance probability $A(x_i', x_i^t)$. As the proposal distributions are symmetric, the term $\frac{g(x_i^t|x_o')}{g(x_i'|x_i^t)}$ equals 1. We further have $\frac{P(\boldsymbol{x}'|\boldsymbol{O})}{P(\boldsymbol{x_t}|\boldsymbol{O})} = \frac{P(\boldsymbol{O}|\boldsymbol{x}')}{P(\boldsymbol{O}|\boldsymbol{x_t})}$, which is the ratio of the two likelihoods. Denote $P(\boldsymbol{O}|\boldsymbol{x})$ as $\exp(\ell(\boldsymbol{x}))$, where $\ell(\boldsymbol{x}) = \log P(\boldsymbol{O}|\boldsymbol{x})$ is the loglikelihood given the parameter vector $\boldsymbol{x}$. The acceptance rate can be written as $A(x_i', x_i^t) = \exp(\Delta\ell)$, where $\Delta\ell = \ell(\boldsymbol{x}') - \ell(\boldsymbol{x_t})$ is the change of loglikelihood after parameter update. Intuitively, if the parameter update does not decrease the loglikelihood (i.e., $\Delta\ell \geq 0$), the candidate parameter is accepted; otherwise (i.e., $\Delta\ell < 0$), the candidate parameter is only accepted with a probability that decreases exponentially with the change of loglikelihood.

To compute the loglikelihood, we assumed Gaussian observation errors in the model. For each set of parameters, we ran the epidemic model throughout the fitting period to generate weekly new reported cases in all neighborhoods. The loglikelihood of observing real-world weekly case data was approximated as $\ell = \sum_\ell \sum_t \log p(y_{lt}^o, \hat{y}_{lt})$, where $p(y_{lt}^o, \hat{y}_{lt})$ is the probability of the observation y_{lt}^o for a Gaussian distribution $\mathcal{N}(\mu = \hat{y}_{lt}, \sigma = 0.2\hat{y}_{lt})$, in which $\hat{y}_{lt}$ is the simulated weekly cases in location l at week t. The summation ran over all neighborhoods and weeks. Other standard deviations of the Gaussian observation errors were tested, and the fitting performance remained similar.

We fit the model to weekly neighborhood-level COVID-19 cases from March 1, 2020 to December 13, 2020 in five phases sequentially. Due to policy changes, model parameters may change over the course of 2020. We therefore split the fitting period into five intervals using dates with major policy changes (Table S2). For each MCMC fitting, we ran 50,000 MCMC steps (i.e., each parameter was updated 50,000 times) and discarded the first 20,000 steps as burn-in. The posterior parameter distributions were estimated using 3,000 samples taken every 10 steps from the last 30,000 steps. The posterior mean obtained from one fitting period was used to initiate the MCMC in the next period. The convergence of MCMC chains were confirmed using the Geweke diagnostic (*65*) (Table S3).

Model selection

To determine the parameters a and b in FOIs, we performed a grid search and model selection. We tested 441 combinations of parameters choosing from $a = \{0,0.1,0.2,\cdots,2\}$ and $b = \{0,0.1,0.2,\cdots,2\}$. For each parameter combination, we independently ran MCMC for the period from March 1, 2020 to June 6, 2020. We focused on this period covering the first wave in NYC because the large variation in crowdedness and dwell time may better differentiate the impacts of these factors on FOIs. The goodness-of-fit was measured using the average loglikelihood in 100 MCMC samples, selected every 10 steps from the last 1,000 updates.

We quantified the uncertainty of parameters using Monte Carlo profile confidence intervals (*38*). We fit the loglikelihood landscape using a cubic spline interpolation and identified the maximum loglikelihood ℓ^* at a^* and b^*. To compute the CI of each parameter, we examined the profile loglikelihood (i.e., the loglikelihood curve when the other parameter was fixed at the maximum-likelihood estimate). The 95% CI was constructed as the set of parameter values with profile loglikelihood higher than $\ell^* - 1.92$ (*38*).

Retrospective forecasts and evaluation metrics

We used the model-data assimilation (M/D/A) framework to generate retrospective forecasts. This framework has been widely used in numerical weather forecasting and infectious disease forecasting (*35*, *66*, *67*). Specifically, we coupled the behavior-driven epidemic model with an efficient data assimilation algorithm, the ensemble adjustment Kalman filter (EAKF) (*40*). In the EAKF, an ensemble of model states was used to represent distributions of model variables and parameters. In our implementation, 500 ensemble members were used. Each week, parameters v_p for five place categories and β_h for baseline transmission were updated so that the model can better fit the most recent observation. The observed model states (weekly confirmed cases) were updated using the Bayes' rule and unobserved model states were updated using their cross-ensemble covariability with observed model states. Forecasts were generated using posterior parameters and variables. We assumed a constant mobility and contact patterns during the short-term forecast horizon. More implementation details and the pseudo-code of the EAKF are provided in Supplementary Materials.

The forecast system generates probabilistic predictions. We evaluated the forecast skills using three metrics: (1) mean absolute percentage error (MAPE), (2) log score, and (3) weighted interval score (WIS) (*43*). For a weekly forecast in a neighborhood, MAPE is defined as $|\bar{y}_p - y_o|/y_o$, where $\bar{y}_p$ is the mean prediction of weekly case and y_o is the observed value. The log score is defined as the logarithmic value (base e) of the fraction of ensemble members that fall within $\pm 25\%$ of the observed value, with a minimal width of ± 50. WIS is a proper scoring rule for quantile forecasts, which was recently used in evaluating probabilistic predictions of infectious diseases (*45*). It converges to the CRPS (continuous ranked probability score) for an increasing number of intervals. We followed the configurations used by the CDC COVID-19 forecast hub and the FluSight challenge, using 23 quantiles (0.01, 0.025, 0.05, 0.1, …, 0.95, 0.975, 0.99) estimated from the 500 ensemble members. The WIS scores were computed using the R function provided at https://github.com/cmu-delphi/covidcast.

The baseline forecast model

We used a metapopulation model to generate forecasts as a baseline. The model dynamics is described by the following equations.

$$\frac{dS_i}{dt} = -S_i \sum_j \frac{\beta \omega(t) M_{ij}(t) I_j}{N_j} - \frac{\beta_h \omega(t) S_i I_i}{N_i}, [4]$$

$$\frac{dE_i}{dt} = S_i \sum_j \frac{\beta \omega(t) M_{ij}(t) I_j}{N_j} + \frac{\beta_h \omega(t) S_i I_i}{N_i} - \frac{E_i}{Z}, [5]$$

$$\frac{dI_i}{dt} = \frac{E_i}{Z} - \frac{I_i}{D}. [6]$$

Here N_i, S_i, E_i, and I_i are the total, susceptible, exposed, and infectious population in neighborhood i; Z and D are the latency and infectious duration; β is the transmission rate in POIs captured by the foot traffic data; β_h is the baseline transmission rate in other uncaptured places; $\omega(t)$ is seasonality term forced by absolute humidity; and $M_{ij}(t)$ represents the fraction of population living in neighborhood i visiting neighborhood j on day t, computed using the same foot traffic data. Note, the baseline model incorporated time-varying cross-neighborhood mobility but did not include the place category-specific FOIs informed by crowdedness and dwell time. This baseline model was coupled with the EAKF to generate retrospective forecasts.

References

1. L. J. Thomas, P. Huang, F. Yin, J. Xu, Z. W. Almquist, J. R. Hipp, C. T. Butts, Geographical patterns of social cohesion drive disparities in early COVID infection hazard. *Proceedings of the National Academy of Sciences* **119**, e2121675119 (2022).

2. B. Hong, B. J. Bonczak, A. Gupta, L. E. Thorpe, C. E. Kontokosta, Exposure density and neighborhood disparities in COVID-19 infection risk. *Proceedings of the National Academy of Sciences* **118**, e2021258118 (2021).

3. H. E. Aschmann, A. R. Riley, R. Chen, Y.-H. Chen, K. Bibbins-Domingo, A. C. Stokes, M. M. Glymour, M. V. Kiang, Dynamics of racial disparities in all-cause mortality during the COVID-19 pandemic. *Proceedings of the National Academy of Sciences* **119**, e2210941119 (2022).

4. M. R. Lamb, S. Kandula, J. Shaman, Differential COVID-19 case positivity in New York City neighborhoods: Socioeconomic factors and mobility. *Influenza and Other Respiratory Viruses* **15**, 209–217 (2021).

5. D. Carrión, E. Colicino, N. F. Pedretti, K. B. Arfer, J. Rush, N. DeFelice, A. C. Just, Neighborhood-level disparities and subway utilization during the COVID-19 pandemic in New York City. *Nat Commun* **12**, 3692 (2021).

6. S. Pei, S. Kandula, J. Cascante Vega, W. Yang, S. Foerster, C. Thompson, J. Baumgartner, S. D. Ahuja, K. Blaney, J. K. Varma, T. Long, J. Shaman, Contact tracing reveals community transmission of COVID-19 in New York City. *Nat Commun* **13**, 6307 (2022).

7. K. Dai, S. Foerster, N. M. Vora, K. Blaney, C. Keeley, L. Hendricks, J. K. Varma, T. Long, J. Shaman, S. Pei, Community transmission of SARS-CoV-2 during the Delta wave in New York City. *BMC Infect Dis* **23**, 753 (2023).

8. K. Koelle, M. A. Martin, R. Antia, B. Lopman, N. E. Dean, The changing epidemiology of SARS-CoV-2. *Science* **375**, 1116–1121 (2022).

9. R. Li, S. Pei, B. Chen, Y. Song, T. Zhang, W. Yang, J. Shaman, Substantial undocumented infection facilitates the rapid dissemination of novel coronavirus (SARS-CoV-2). *Science* **368**, 489–493 (2020).

10. M. Chinazzi, J. T. Davis, M. Ajelli, C. Gioannini, M. Litvinova, S. Merler, A. P. y Piontti, K. Mu, L. Rossi, K. Sun, C. Viboud, X. Xiong, H. Yu, M. E. Halloran, I. M. Longini, A. Vespignani, The effect of travel restrictions on the spread of the 2019 novel coronavirus (COVID-19) outbreak. *Science* **368**, 395–400 (2020).

11. J. S. Jia, X. Lu, Y. Yuan, G. Xu, J. Jia, N. A. Christakis, Population flow drives spatio-temporal distribution of COVID-19 in China. *Nature* **582**, 389–394 (2020).

12. M. U. G. Kraemer, C.-H. Yang, B. Gutierrez, C.-H. Wu, B. Klein, D. M. Pigott, O. C.-19 D. W. Group†, L. du Plessis, N. R. Faria, R. Li, W. P. Hanage, J. S. Brownstein, M. Layan, A. Vespignani, H. Tian, C. Dye, O. G. Pybus, S. V. Scarpino, The effect of human mobility and control measures on the COVID-19 epidemic in China. *Science* **368**, 493–497 (2020).

13. S. Lai, N. W. Ruktanonchai, L. Zhou, O. Prosper, W. Luo, J. R. Floyd, A. Wesolowski, M. Santillana, C. Zhang, X. Du, H. Yu, A. J. Tatem, Effect of non-pharmaceutical interventions to contain COVID-19 in China. *Nature* **585**, 410–413 (2020).

14. H. Tian, Y. Liu, Y. Li, C.-H. Wu, B. Chen, M. U. G. Kraemer, B. Li, J. Cai, B. Xu, Q. Yang, B. Wang, P. Yang, Y. Cui, Y. Song, P. Zheng, Q. Wang, O. N. Bjornstad, R. Yang, B. T. Grenfell, O. G. Pybus, C. Dye, An investigation of transmission control measures during the first 50 days of the COVID-19 epidemic in China. *Science* **368**, 638–642 (2020).

15. C. O. Buckee, S. Balsari, J. Chan, M. Crosas, F. Dominici, U. Gasser, Y. H. Grad, B. Grenfell, M. E. Halloran, M. U. G. Kraemer, M. Lipsitch, C. J. E. Metcalf, L. A. Meyers, T. A. Perkins, M. Santillana, S. V. Scarpino, C. Viboud, A. Wesolowski, A. Schroeder, Aggregated mobility data could help fight COVID-19. *Science* **368**, 145–146 (2020).

16. S. M. Kissler, N. Kishore, M. Prabhu, D. Goffman, Y. Beilin, R. Landau, C. Gyamfi-Bannerman, B. T. Bateman, J. Snyder, A. S. Razavi, D. Katz, J. Gal, A. Bianco, J. Stone, D. Larremore, C. O. Buckee, Y. H. Grad, Reductions in commuting mobility correlate with geographic differences in SARS-CoV-2 prevalence in New York City. *Nat Commun* **11**, 4674 (2020).

17. P. Nouvellet, S. Bhatia, A. Cori, K. E. C. Ainslie, M. Baguelin, S. Bhatt, A. Boonyasiri, N. F. Brazeau, L. Cattarino, L. V. Cooper, H. Coupland, Z. M. Cucunuba, G. Cuomo-Dannenburg, A. Dighe, B. A. Djaafara, I. Dorigatti, O. D. Eales, S. L. van Elsland, F. F. Nascimento, R. G. FitzJohn, K. A. M. Gaythorpe, L. Geidelberg, W. D. Green, A. Hamlet, K. Hauck, W. Hinsley, N. Imai, B. Jeffrey, E. Knock, D. J. Laydon, J. A. Lees, T. Mangal, T. A. Mellan, G. Nedjati-Gilani, K. V. Parag, M. Pons-Salort, M. Ragonnet-Cronin, S. Riley, H. J. T. Unwin, R. Verity, M. A. C. Vollmer, E. Volz, P. G. T. Walker, C. E. Walters, H. Wang, O. J. Watson, C. Whittaker, L. K. Whittles, X. Xi, N. M. Ferguson, C. A. Donnelly, Reduction in mobility and COVID-19 transmission. *Nat Commun* **12**, 1090 (2021).

18. H. S. Badr, H. Du, M. Marshall, E. Dong, M. M. Squire, L. M. Gardner, Association between mobility patterns and COVID-19 transmission in the USA: a mathematical modelling study. *The Lancet Infectious Diseases* **20**, 1247–1254 (2020).

19. H. S. Badr, L. M. Gardner, Limitations of using mobile phone data to model COVID-19 transmission in the USA. *The Lancet Infectious Diseases* **21**, E113 (2020).

20. N. Kishore, A. R. Taylor, P. E. Jacob, N. Vembar, T. Cohen, C. O. Buckee, N. A. Menzies, Evaluating the reliability of mobility metrics from aggregated mobile phone data as proxies for SARS-CoV-2 transmission in the USA: a population-based study. *The Lancet Digital Health* **4**, e27–e36 (2022).

21. S. Chang, E. Pierson, P. W. Koh, J. Gerardin, B. Redbird, D. Grusky, J. Leskovec, Mobility network models of COVID-19 explain inequities and inform reopening. *Nature* **589**, 82–87 (2021).

22. X. Hou, S. Gao, Q. Li, Y. Kang, N. Chen, K. Chen, J. Rao, J. S. Ellenberg, J. A. Patz, Intracounty modeling of COVID-19 infection with human mobility: Assessing spatial heterogeneity with business traffic, age, and race. *Proceedings of the National Academy of Sciences* **118**, e2020524118 (2021).

23. S. Chang, M. L. Wilson, B. Lewis, Z. Mehrab, K. K. Dudakiya, E. Pierson, P. W. Koh, J. Gerardin, B. Redbird, D. Grusky, M. Marathe, J. Leskovec, "Supporting COVID-19 Policy Response with Large-scale Mobility-based Modeling" in *Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining* (Association for Computing Machinery, New York, NY, USA, 2021; https://dl.acm.org/doi/10.1145/3447548.3467182)*KDD '21*, pp. 2632–2642.

24. S. Venkatramanan, B. Lewis, J. Chen, D. Higdon, A. Vullikanti, M. Marathe, Using data-driven agent-based models for forecasting emerging infectious diseases. *Epidemics* **22**, 43–49 (2018).

25. J. J. Grefenstette, S. T. Brown, R. Rosenfeld, J. DePasse, N. T. Stone, P. C. Cooley, W. D. Wheaton, A. Fyshe, D. D. Galloway, A. Sriram, H. Guclu, T. Abraham, D. S. Burke, FRED (A Framework for Reconstructing Epidemic Dynamics): an open-source software system for modeling infectious diseases and control strategies using census-based populations. *BMC Public Health* **13**, 940 (2013).

26. Z. Susswein, E. C. Rest, S. Bansal, Disentangling the rhythms of human activity in the built environment for airborne transmission risk: An analysis of large-scale mobility data. *eLife* **12**, e80466 (2023).

27. J. Zhang, S. Tan, C. Peng, X. Xu, M. Wang, W. Lu, Y. Wu, B. Sai, M. Cai, A. G. Kummer, Z. Chen, J. Zou, W. Li, W. Zheng, Y. Liang, Y. Zhao, A. Vespignani, M. Ajelli, X. Lu, H. Yu, Heterogeneous changes in mobility in response to the SARS-CoV-2 Omicron BA.2 outbreak in Shanghai. *Proceedings of the National Academy of Sciences* **120**, e2306710120 (2023).

28. W. Li, Q. Wang, Y. Liu, M. L. Small, J. Gao, A spatiotemporal decay model of human mobility when facing large-scale crises. *Proceedings of the National Academy of Sciences* **119**, e2203042119 (2022).

29. SafeGraph, Places Data Curated for Accurate Geospatial Analytics | SafeGraph. https://safegraph.com/.

30. NAICS, North American Industry Classification System (NAICS) U.S. Census Bureau. https://www.census.gov/naics/.

31. NYC Planning, NYC geography of jobs. https://www.nyc.gov/assets/planning/download/pdf/planning-level/housing-economy/nyc-geography-jobs2-1019.pdf.

32. B. Rader, S. V. Scarpino, A. Nande, A. L. Hill, B. Adlam, R. C. Reiner, D. M. Pigott, B. Gutierrez, A. E. Zarebski, M. Shrestha, J. S. Brownstein, M. C. Castro, C. Dye, H. Tian, O. G. Pybus, M. U. G. Kraemer, Crowding and the shape of COVID-19 epidemics. *Nat Med* **26**, 1829–1834 (2020).

33. Y. Ma, S. Pei, J. Shaman, R. Dubrow, K. Chen, Role of meteorological factors in the transmission of SARS-CoV-2 in the United States. *Nat Commun* **12**, 3602 (2021).

34. J. Shaman, M. Kohn, Absolute humidity modulates influenza survival, transmission, and seasonality. *Proceedings of the National Academy of Sciences* **106**, 3243–3248 (2009).

35. J. Shaman, A. Karspeck, Forecasting seasonal outbreaks of influenza. *Proceedings of the National Academy of Sciences* **109**, 20425–20430 (2012).

36. W. Yang, S. K. Greene, E. R. Peterson, W. Li, R. Mathes, L. Graf, R. Lall, S. Hughes, J. Wang, A. Fine, Epidemiological characteristics of the B.1.526 SARS-CoV-2 variant. *Science Advances* **8**, eabm0300.

37. S. Brooks, A. Gelman, G. Jones, X.-L. Meng, *Handbook of Markov Chain Monte Carlo* (CRC Press, 2011).

38. E. L. Ionides, C. Breto, J. Park, R. A. Smith, A. A. King, Monte Carlo profile confidence intervals for dynamic systems. *J. R. Soc. Interface.* **14**, 20170126 (2017).

39. D. He, E. L. Ionides, A. A. King, Plug-and-play inference for disease dynamics: measles in large and small populations as a case study. *Journal of The Royal Society Interface* **7**, 271–283 (2010).

40. J. L. Anderson, An Ensemble Adjustment Kalman Filter for Data Assimilation. *Mon. Wea. Rev.* **129**, 2884–2903 (2001).

41. S. Pei, T. K. Yamana, S. Kandula, M. Galanti, J. Shaman, Burden and characteristics of COVID-19 in the United States during 2020. *Nature* **598**, 338–341 (2021).

42. S. Pei, S. Kandula, J. Shaman, Differential effects of intervention timing on COVID-19 spread in the United States. *Science Advances* **6**, eabd6370 (2020).

43. J. Bracher, E. L. Ray, T. Gneiting, N. G. Reich, Evaluating epidemic forecasts in an interval format. *PLOS Computational Biology* **17**, e1008618 (2021).

44. N. G. Reich, L. C. Brooks, S. J. Fox, S. Kandula, C. J. McGowan, E. Moore, D. Osthus, E. L. Ray, A. Tushar, T. K. Yamana, M. Biggerstaff, M. A. Johansson, R. Rosenfeld, J. Shaman, A collaborative multiyear, multimodel assessment of seasonal influenza forecasting in the United States. *Proceedings of the National Academy of Sciences* **116**, 3146–3154 (2019).

45. E. Y. Cramer, E. L. Ray, V. K. Lopez, J. Bracher, A. Brennen, A. J. Castro Rivadeneira, A. Gerding, T. Gneiting, K. H. House, Y. Huang, D. Jayawardena, A. H. Kanji, A. Khandelwal, K. Le, A. Mühlemann, J. Niemi, A. Shah, A. Stark, Y. Wang, N. Wattanachit, M. W. Zorn, Y. Gu, S. Jain, N. Bannur, A. Deva, M. Kulkarni, S. Merugu, A. Raval, S. Shingi, A. Tiwari, J. White, N. F. Abernethy, S. Woody, M. Dahan, S. Fox, K. Gaither, M. Lachmann, L. A. Meyers, J. G. Scott, M. Tec, A. Srivastava, G. E. George, J. C. Cegan, I. D. Dettwiller, W. P. England, M. W. Farthing, R. H. Hunter, B. Lafferty, I. Linkov, M. L. Mayo, M. D. Parno, M. A. Rowland, B. D. Trump, Y. Zhang-James, S. Chen, S. V. Faraone, J. Hess, C. P. Morley, A. Salekin, D. Wang, S. M. Corsetti, T. M. Baer, M. C. Eisenberg, K. Falb, Y. Huang, E. T. Martin, E. McCauley, R. L. Myers, T. Schwarz, D. Sheldon, G. C. Gibson, R. Yu, L. Gao, Y. Ma, D. Wu, X. Yan, X. Jin, Y.-X. Wang, Y. Chen, L. Guo, Y. Zhao, Q. Gu, J. Chen, L. Wang, P. Xu, W. Zhang, D. Zou, H. Biegel, J. Lega, S. McConnell, V. P. Nagraj, S. L. Guertin, C. Hulme-Lowe, S. D. Turner, Y. Shi, X. Ban, R. Walraven, Q.-J. Hong, S. Kong, A. van de Walle, J. A. Turtle, M. Ben-Nun, S. Riley, P. Riley, U. Koyluoglu, D. DesRoches, P. Forli, B. Hamory, C. Kyriakides, H. Leis, J. Milliken, M. Moloney, J. Morgan, N. Nirgudkar, G. Ozcan, N. Piwonka, M. Ravi, C. Schrader, E. Shakhnovich, D. Siegel, R. Spatz, C. Stiefeling, B. Wilkinson, A. Wong, S. Cavany, G. España, S. Moore, R.

Oidtman, A. Perkins, D. Kraus, A. Kraus, Z. Gao, J. Bian, W. Cao, J. Lavista Ferres, C. Li, T.-Y. Liu, X. Xie, S. Zhang, S. Zheng, A. Vespignani, M. Chinazzi, J. T. Davis, K. Mu, A. Pastore y Piontti, X. Xiong, A. Zheng, J. Baek, V. Farias, A. Georgescu, R. Levi, D. Sinha, J. Wilde, G. Perakis, M. A. Bennouna, D. Nze-Ndong, D. Singhvi, I. Spantidakis, L. Thayaparan, A. Tsiourvas, A. Sarker, A. Jadbabaie, D. Shah, N. Della Penna, L. A. Celi, S. Sundar, R. Wolfinger, D. Osthus, L. Castro, G. Fairchild, I. Michaud, D. Karlen, M. Kinsey, L. C. Mullany, K. Rainwater-Lovett, L. Shin, K. Tallaksen, S. Wilson, E. C. Lee, J. Dent, K. H. Grantz, A. L. Hill, J. Kaminsky, K. Kaminsky, L. T. Keegan, S. A. Lauer, J. C. Lemaitre, J. Lessler, H. R. Meredith, J. Perez-Saez, S. Shah, C. P. Smith, S. A. Truelove, J. Wills, M. Marshall, L. Gardner, K. Nixon, J. C. Burant, L. Wang, L. Gao, Z. Gu, M. Kim, X. Li, G. Wang, Y. Wang, S. Yu, R. C. Reiner, R. Barber, E. Gakidou, S. I. Hay, S. Lim, C. Murray, D. Pigott, H. L. Gurung, P. Baccam, S. A. Stage, B. T. Suchoski, B. A. Prakash, B. Adhikari, J. Cui, A. Rodríguez, A. Tabassum, J. Xie, P. Keskinocak, J. Asplund, A. Baxter, B. E. Oruc, N. Serban, S. O. Arik, M. Dusenberry, A. Epshteyn, E. Kanal, L. T. Le, C.-L. Li, T. Pfister, D. Sava, R. Sinha, T. Tsai, N. Yoder, J. Yoon, L. Zhang, S. Abbott, N. I. Bosse, S. Funk, J. Hellewell, S. R. Meakin, K. Sherratt, M. Zhou, R. Kalantari, T. K. Yamana, S. Pei, J. Shaman, M. L. Li, D. Bertsimas, O. Skali Lami, S. Soni, H. Tazi Bouardi, T. Ayer, M. Adee, J. Chhatwal, O. O. Dalgic, M. A. Ladd, B. P. Linas, P. Mueller, J. Xiao, Y. Wang, Q. Wang, S. Xie, D. Zeng, A. Green, J. Bien, L. Brooks, A. J. Hu, M. Jahja, D. McDonald, B. Narasimhan, C. Politsch, S. Rajanala, A. Rumack, N. Simon, R. J. Tibshirani, R. Tibshirani, V. Ventura, L. Wasserman, E. B. O'Dea, J. M. Drake, R. Pagano, Q. T. Tran, L. S. T. Ho, H. Huynh, J. W. Walker, R. B. Slayton, M. A. Johansson, M. Biggerstaff, N. G. Reich, Evaluation of individual and ensemble probabilistic forecasts of COVID-19 mortality in the United States. *Proceedings of the National Academy of Sciences* **119**, e2113561119 (2022).

46. Y. Yang, Y. Li, K. Kral, N. Hupert, T. Dogan, Urban design attributes and resilience: COVID-19 evidence from New York City. *Buildings and Cities* **2**, 618–636 (2021).

47. C. P. Tribby, C. Hartmann, COVID-19 Cases and the Built Environment: Initial Evidence from New York City. *The Professional Geographer* **73**, 365–376 (2021).

48. K. Cranmer, J. Brehmer, G. Louppe, The frontier of simulation-based inference. *PNAS* **117**, 30055–30062 (2020).

49. J. Bedson, L. A. Skrip, D. Pedi, S. Abramowitz, S. Carter, M. F. Jalloh, S. Funk, N. Gobat, T. Giles-Vernick, G. Chowell, J. R. de Almeida, R. Elessawi, S. V. Scarpino, R. A. Hammond, S. Briand, J. M. Epstein, L. Hébert-Dufresne, B. M. Althouse, A review and agenda for integrated disease models including social and behavioural factors. *Nat Hum Behav* **5**, 834–846 (2021).

50. J. J. V. Bavel, K. Baicker, P. S. Boggio, V. Capraro, A. Cichocka, M. Cikara, M. J. Crockett, A. J. Crum, K. M. Douglas, J. N. Druckman, J. Drury, O. Dube, N. Ellemers, E. J. Finkel, J. H. Fowler, M. Gelfand, S. Han, S. A. Haslam, J. Jetten, S. Kitayama, D. Mobbs, L. E. Napper, D. J. Packer, G. Pennycook, E. Peters, R. E. Petty, D. G. Rand, S. D. Reicher, S. Schnall, A. Shariff, L. J. Skitka, S. S. Smith, C. R. Sunstein, N. Tabri, J. A. Tucker, S. van der Linden, P. van Lange, K. A. Weeden, M. J. A. Wohl, J. Zaki, S. R. Zion, R. Willer, Using social and behavioural science to support COVID-19 pandemic response. *Nat Hum Behav* **4**, 460–471 (2020).

51. E. P. Fenichel, C. Castillo-Chavez, M. G. Ceddia, G. Chowell, P. A. G. Parra, G. J. Hickling, G. Holloway, R. Horan, B. Morin, C. Perrings, M. Springborn, L. Velazquez, C. Villalobos, Adaptive human behavior in epidemiological models. *PNAS* **108**, 6306–6311 (2011).

52. Z. Wang, M. A. Andrews, Z.-X. Wu, L. Wang, C. T. Bauch, Coupled disease–behavior dynamics on complex networks: A review. *Physics of Life Reviews* **15**, 1–29 (2015).

53. J. S. Weitz, S. W. Park, C. Eksin, J. Dushoff, Awareness-driven behavior changes can shift the shape of epidemics away from peaks and toward plateaus, shoulders, and oscillations. *PNAS* **117**, 32764–32771 (2020).

54. C. Granell, S. Gómez, A. Arenas, Dynamical Interplay between Awareness and Epidemic Spreading in Multiplex Networks. *Phys. Rev. Lett.* **111**, 128701 (2013).

55. F. Verelst, L. Willem, P. Beutels, Behavioural change models for infectious disease transmission: a systematic review (2010–2015). *Journal of The Royal Society Interface* **13**, 20160820 (2016).

56. S. Funk, M. Salathé, V. A. A. Jansen, Modelling the influence of human behaviour on the spread of infectious diseases: a review. *Journal of The Royal Society Interface* **7**, 1247–1256 (2010).

57. K. Ruggeri, F. Stock, S. A. Haslam, V. Capraro, P. Boggio, N. Ellemers, A. Cichocka, K. M. Douglas, D. G. Rand, S. van der Linden, M. Cikara, E. J. Finkel, J. N. Druckman, M. J. A. Wohl, R. E. Petty, J. A. Tucker, A. Shariff, M. Gelfand, D. Packer, J. Jetten, P. A. M. Van Lange, G. Pennycook, E. Peters, K. Baicker, A. Crum, K. A. Weeden, L. Napper, N. Tabri, J. Zaki, L. Skitka, S. Kitayama, D. Mobbs, C. R. Sunstein, S. Ashcroft-Jones, A. L. Todsen, A. Hajian, S. Verra, V. Buehler, M. Friedemann, M. Hecht, R. S. Mobarak, R. Karakasheva, M. R. Tünte, S. K. Yeung, R. S. Rosenbaum, Ž. Lep, Y. Yamada, S. T. J. Hudson, L. Macchia, I. Soboleva, E. Dimant, S. J. Geiger, H. Jarke, T. Wingen, J. B. Berkessel, S. Mareva, L. McGill, F. Papa, B. Većkalov, Z. Afif, E. K. Buabang, M. Landman, F. Tavera, J. L. Andrews, A. Bursalıoğlu, Z. Zupan, L. Wagner, J. Navajas, M. Vranka, D. Kasdan, P. Chen, K. R. Hudson, L. M. Novak, P. Teas, N. R. Rachev, M. M. Galizzi, K. L. Milkman, M. Petrović, J. J. Van Bavel, R. Willer, A synthesis of evidence for policy from behavioural science during COVID-19. *Nature* **625**, 134–147 (2024).

58. K. Sun, S. Tempia, J. Kleynhans, A. von Gottberg, M. L. McMorrow, N. Wolter, J. N. Bhiman, J. Moyes, M. Carrim, N. A. Martinson, K. Kahn, L. Lebina, J. D. du Toit, T. Mkhencele, C. Viboud, C. Cohen, Rapidly shifting immunologic landscape and severity of SARS-CoV-2 in the Omicron era in South Africa. *Nat Commun* **14**, 246 (2023).

59. P. V. Markov, M. Ghafari, M. Beer, K. Lythgoe, P. Simmonds, N. I. Stilianakis, A. Katzourakis, The evolution of SARS-CoV-2. *Nat Rev Microbiol* **21**, 361–379 (2023).

60. N. D. Rochman, Y. I. Wolf, G. Faure, P. Mutz, F. Zhang, E. V. Koonin, Ongoing global and regional adaptive evolution of SARS-CoV-2. *PNAS* **118** (2021).

61. NLDAS, NLDAS: Project Goals | LDAS. https://ldas.gsfc.nasa.gov/nldas.

62. V. J. Hall, S. Foulkes, A. Charlett, A. Atti, E. J. M. Monk, R. Simmons, E. Wellington, M. J. Cole, A. Saei, B. Oguti, K. Munro, S. Wallace, P. D. Kirwan, M. Shrotri, A. Vusirikala, S. Rokadiya, M. Kall, M. Zambon, M. Ramsay, T. Brooks, C. S. Brown, M. A. Chand, S.

Hopkins, N. Andrews, A. Atti, H. Aziz, T. Brooks, C. S. Brown, D. Camero, C. Carr, M. A. Chand, A. Charlett, H. Crawford, M. Cole, J. Conneely, S. D'Arcangelo, J. Ellis, S. Evans, S. Foulkes, N. Gillson, R. Gopal, L. Hall, V. J. Hall, P. Harrington, S. Hopkins, J. Hewson, K. Hoschler, D. Ironmonger, J. Islam, M. Kall, I. Karagiannis, O. Kay, J. Khawam, E. King, P. Kirwan, R. Kyffin, A. Lackenby, M. Lattimore, E. Linley, J. Lopez-Bernal, L. Mabey, R. McGregor, S. Miah, E. J. M. Monk, K. Munro, Z. Naheed, A. Nissr, A. M. O'Connell, B. Oguti, H. Okafor, S. Organ, J. Osbourne, A. Otter, M. Patel, S. Platt, D. Pople, K. Potts, M. Ramsay, J. Robotham, S. Rokadiya, C. Rowe, A. Saei, G. Sebbage, A. Semper, M. Shrotri, R. Simmons, A. Soriano, P. Staves, S. Taylor, A. Taylor, A. Tengbe, S. Tonge, A. Vusirikala, S. Wallace, E. Wellington, M. Zambon, D. Corrigan, M. Sartaj, L. Cromey, S. Campbell, K. Braithwaite, L. Price, L. Haahr, S. Stewart, E. D. Lacey, L. Partridge, G. Stevens, Y. Ellis, H. Hodgson, C. Norman, B. Larru, S. Mcwilliam, S. Winchester, P. Cieciwa, A. Pai, C. Loughrey, A. Watt, F. Adair, A. Hawkins, A. Grant, R. Temple-Purcell, J. Howard, N. Slawson, C. Subudhi, S. Davies, A. Bexley, R. Penn, N. Wong, G. Boyd, A. Rajgopal, A. Arenas-Pinto, R. Matthews, A. Whileman, R. Laugharne, J. Ledger, T. Barnes, C. Jones, D. Botes, N. Chitalia, S. Akhtar, G. Harrison, S. Horne, N. Walker, K. Agwuh, V. Maxwell, J. Graves, S. Williams, A. O'Kelly, P. Ridley, A. Cowley, H. Johnstone, P. Swift, J. Democratis, M. Meda, C. Callens, S. Beazer, S. Hams, V. Irvine, B. Chandrasekaran, C. Forsyth, J. Radmore, C. Thomas, K. Brown, S. Roberts, P. Burns, K. Gajee, T. M. Byrne, F. Sanderson, S. Knight, E. Macnaughton, B. J. L. Burton, H. Smith, R. Chaudhuri, K. Hollinshead, R. J. Shorten, A. Swan, R. J. Shorten, C. Favager, J. Murira, S. Baillon, S. Hamer, K. Gantert, J. Russell, D. Brennan, A. Dave, A. Chawla, F. Westell, D. Adeboyeku, P. Papineni, C. Pegg, M. Williams, S. Ahmad, S. Ingram, C. Gabriel, K. Pagget, P. Cieciwa, G. Maloney, J. Ashcroft, I. D. Rosario, R. Crosby-Nwaobi, C. Reeks, S. Fowler, L. Prentice, M. Spears, G. McKerron, K. McLelland-Brooks, J. Anderson, S. Donaldson, K. Templeton, L. Coke, N. Elumogo, J. Elliott, D. Padgett, M. Mirfenderesky, A. Cross, J. Price, S. Joyce, I. Sinanovic, M. Howard, T. Lewis, P. Cowling, D. Potoczna, S. Brand, L. Sheridan, B. Wadams, A. Lloyd, J. Mouland, J. Giles, G. Pottinger, H. Coles, M. Joseph, M. Lee, S. Orr, H. Chenoweth, C. Auckland, R. Lear, T. Mahungu, A. Rodger, K. Penny-Thomas, S. Pai, J. Zamikula, E. Smith, S. Stone, E. Boldock, D. Howcroft, C. Thompson, M. Aga, P. Domingos, S. Gormley, C. Kerrison, L. Marsh, S. Tazzyman, L. Allsop, S. Ambalkar, M. Beekes, S. Jose, J. Tomlinson, A. Jones, C. Price, J. Pepperell, M. Schultz, J. Day, A. Boulos, E. Defever, D. McCracken, K. Brown, K. Gray, A. Houston, T. Planche, R. P. Jones, D. Wycherley, S. Bennett, J. Marrs, K. Nimako, B. Stewart, N. Kalakonda, S. Khanduri, A. Ashby, M. Holden, N. Mahabir, J. Harwood, B. Payne, K. Court, N. Staines, R. Longfellow, M. E. Green, L. E. Hughes, M. Halkes, P. Mercer, A. Roebuck, E. Wilson-Davies, L. Gallego, R. Lazarus, N. Aldridge, L. Berry, F. Game, T. Reynolds, C. Holmes, M. Wiselka, A. Higham, M. Booth, C. Duff, J. Alderton, H. Jory, E. Virgilio, T. Chin, M. Z. Qazzafi, A. M. Moody, R. Tilley, T. Donaghy, K. Shipman, R. Sierra, N. Jones, G. Mills, D. Harvey, Y. W. J. Huang, J. Birch, L. Robinson, S. Board, A. Broadley, C. Laven, N. Todd, D. W. Eyre, K. Jeffery, S. Dunachie, C. Duncan, P. Klenerman, L. Turtle, T. D. Silva, H. Baxendale, J. L. Heeney, SARS-CoV-2 infection rates of antibody-positive compared with antibody-negative health-care workers in England: a large, multicentre, prospective cohort study (SIREN). *The Lancet* **397**, 1459–1469 (2021).

63. W. Yang, S. Kandula, M. Huynh, S. K. Greene, G. V. Wye, W. Li, H. T. Chan, E. McGibbon, A. Yeung, D. Olson, A. Fine, J. Shaman, Estimating the infection-fatality risk of SARS-CoV-2 in New York City during the spring 2020 pandemic wave: a model-based analysis. *The Lancet Infectious Diseases* **21**, 203–212 (2021).

64. S. Pei, J. Shaman, Initial Simulation of SARS-CoV2 Spread and Intervention Effects in the Continental US. *medRxiv*, 2020.03.21.20040303 (2020).

65. J. Geweke, "Evaluating the Accuracy of Sampling-Based Approaches to the Calculation of Posterior Moments" in *In Bayesian Statistics* (University Press, 1992), pp. 169–193.

66. W. Yang, A. Karspeck, J. Shaman, Comparison of Filtering Methods for the Modeling and Retrospective Forecasting of Influenza Epidemics. *PLOS Computational Biology* **10**, e1003583 (2014).

67. S. Pei, S. Kandula, W. Yang, J. Shaman, Forecasting the spatial transmission of influenza in the United States. *PNAS* **115**, 2752–2757 (2018).

Acknowledgments

We thank SafeGraph for sharing the foot traffic data and NYC DOHMH for providing real-world neighborhood-level COVID-19 case data. We thank Hilary Parton, Ellen Lee, Lauren Firestein, and Shama Ahuja at NYC DOHMH for comments and feedback.

Competing interests:

JS and Columbia University disclose partial ownership of SK Analytics. JS discloses consulting for BNI. All other authors declare they have no competing interests.

Figures and Tables

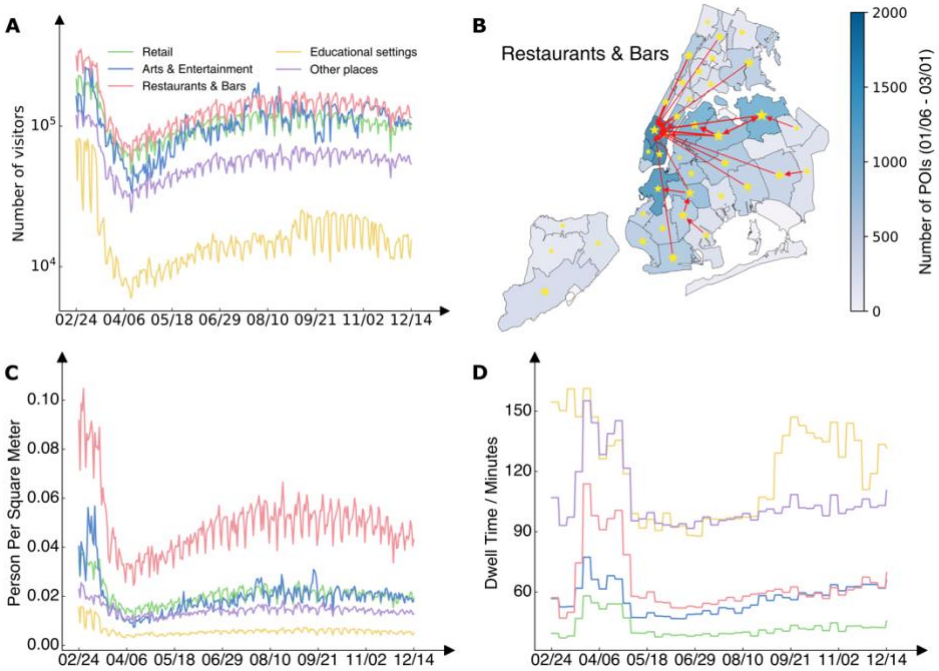

Fig. 1. Mobility and contact patterns in different place categories. (**A**), Daily visitor counts to five place categories (restaurants & bars, retail, arts & entertainment, educational settings, and others) in NYC during 2020, as recorded in the foot traffic data. (**B**), Geographical distribution of restaurants & bars across NYC neighborhoods (color). Stars and arrows highlight mobility links with over 1000 visitors per day from January 6, 2020 to March 1, 2020. Stars indicate self-links where residents visited restaurants & bars within their own neighborhoods. (**C**), Daily average crowdedness (daily visitor counts per square meter) at POIs for five place categories. (**D**), Daily average dwell time (minutes) at POIs for five place categories.

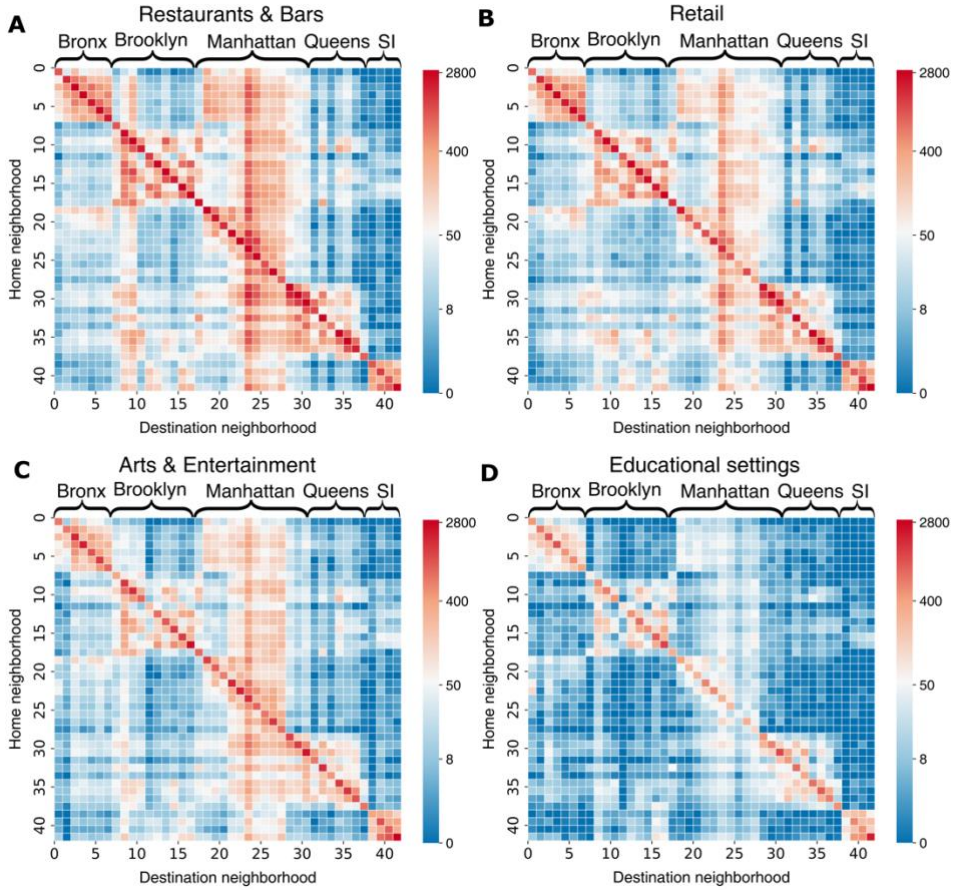

Fig. 2. Mobility patterns across 42 NYC neighborhoods in four place categories. Daily average visitor counts (in log scale) from home neighborhoods (y-axis) to destination neighborhoods (x-axis) in restaurants & bars (**A**), retail (**B**), arts & entertainment (**C**), and educational settings (**D**). The five boroughs of NYC (the Bronx, Brooklyn, Manhattan, Queens, and Staten Island) are indicated on top of each heatmap. Foot traffic data from January 6, 2020 to March 1, 2020 were used, representing the period before the implementation of governmental interventions.

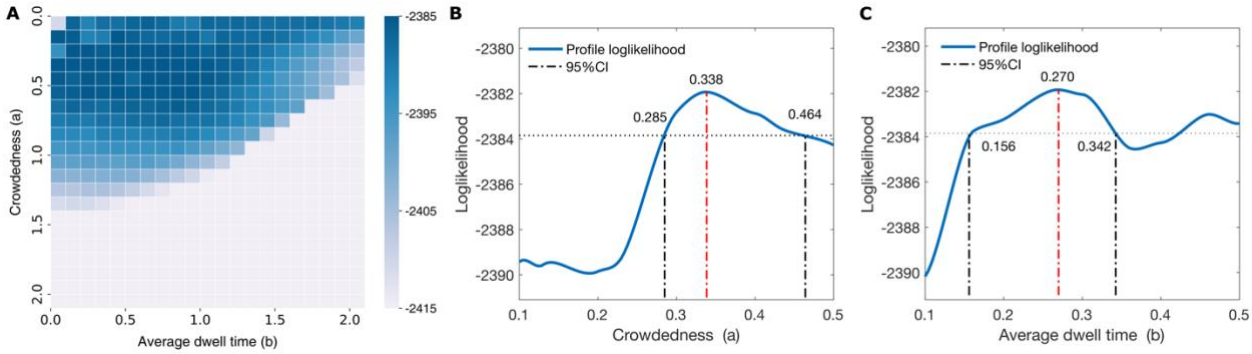

Fig. 3. Estimating the impacts of crowdedness and dwell time on force of infection. (**A**), The loglikelihood landscape for different combinations of parameters a (the power exponent for crowdedness) and b (the power exponent for average dwell time). The loglikelihood for each parameter combination was obtained by averaging the loglikelihood values over the last 100 MCMC samples (selected every 10 steps from the last 1,000 samples in the MCMC chain). (**B**), The profile loglikelihood for parameter a near the maximum of the loglikelihood surface. The 95% CI is constructed as the set of parameter values with profile loglikelihood higher than $\ell^* - 1.92$, where ℓ^* is the maximum loglikelihood. The vertical red dash line shows the estimate of a and the vertical black dash lines show the estimated 95% CI. (**C**), The profile loglikelihood for parameter b near the maximum of the loglikelihood surface.

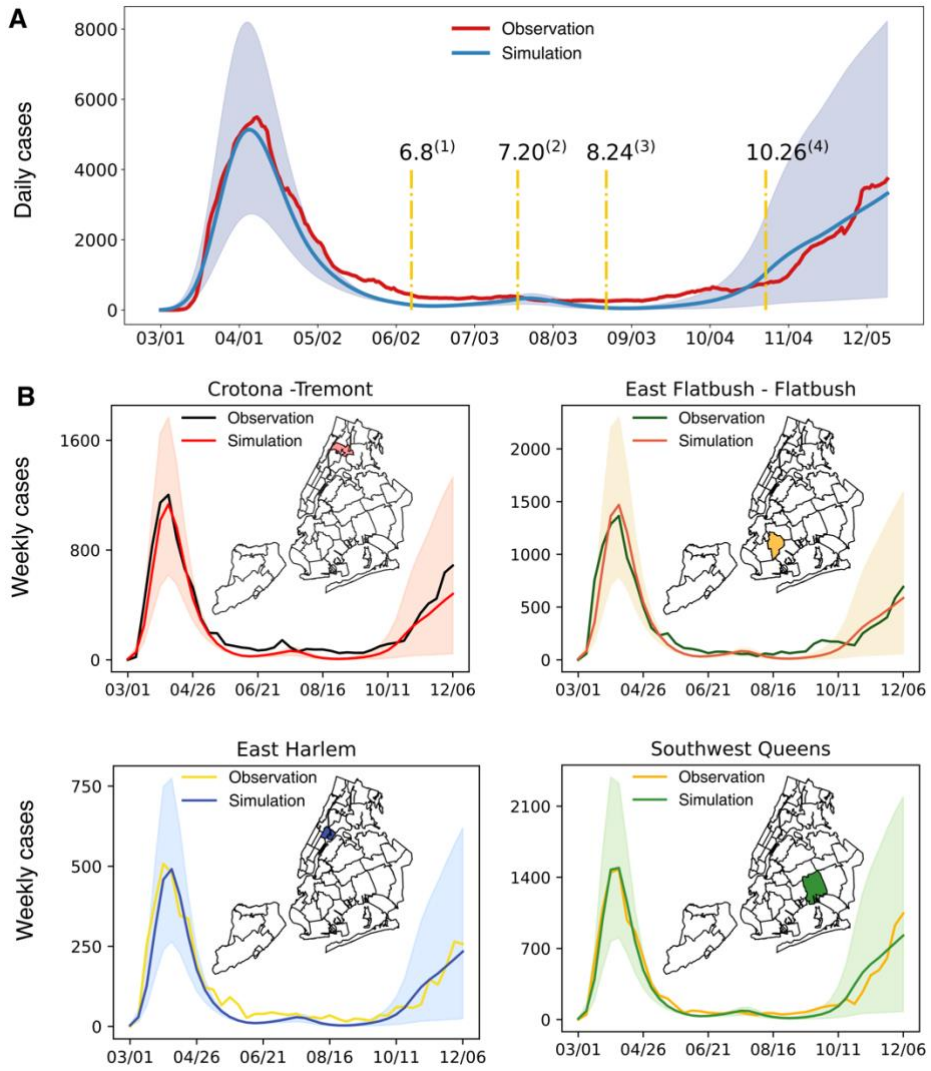

Fig. 4. Model fitting to neighborhood-level COVID-19 case data. (**A**), Simulations using model parameters estimated for five separate periods, as indicated by the yellow vertical lines corresponding to dates with major policy changes (Table S2). Simulated cases were aggregated to the city level and are compared with the daily confirmed cases in NYC (red line). The shaded blue area represents the 95% confidence interval, obtained from 500 independent simulations. (**B**), Simulations in four representative neighborhoods in the Bronx (upper left), Brooklyn (upper right), Manhattan (lower left), and Queens (lower right). Maps display the geographical locations of these neighborhoods.

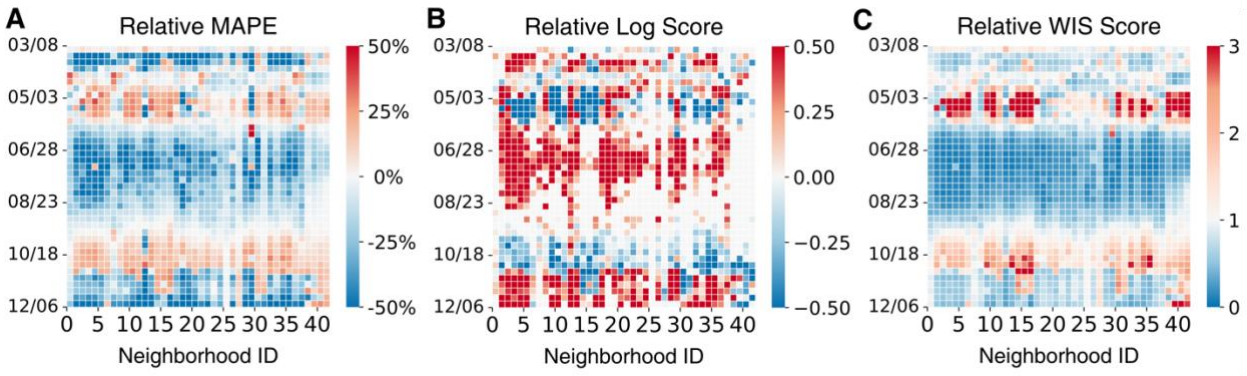

Fig. 5. One-week ahead retrospective forecasts for neighborhood-level COVID-19 cases. The behavior-driven forecasts are compared with a baseline model that does not distinguish mobility and contact patterns in different place categories. From left to right, we present (**A**) the relative mean absolute percentage error (MAPE) (MAPEs of the behavior-driven model minus those of the baseline, with blue indicating better forecasts), (**B**) relative log score (log scores of the behavior-driven model minus those of the baseline, with red indicating better forecasts), and (**C**) relative weighted interval score (WIS) (the *ratio* of the WIS scores of the behavior-driven model to those of the baseline, with blue indicating better forecasts) for all 42 neighborhoods from March 8, 2020 to December 13, 2020.